VAA-Weekly-Progress

03/04-03/11

Context

- Discussed experiments of species similarity metrics
- Finalized keypoint definitions and wanted to begin labeling

Goals

- Everyone label 40 (33 + 7 overlap) images for Visible Definitions
- Shaan 1- 33, 40-47
- Claire 34 - 66, 73-80
- Josh 67 - 100, 107-114
- Zian 101 - 134, 141-148,
- Medha 135 - 167, 174-181
- Parth 168 - 200, 6-13
- Professor: 1-5, 34-39, 67-72, 101-106, 135-140, 168-173

- Setup label studio for annotating antelope images (Shaan)
- Run Dino again with cropped bounding boxes and add Stable Diffusion features (Josh)
- Running centroid variation (Medha), limb ratio (Zian), dino (Josh) (without side-view images)
- Look into downstream task metrics (Claire)
- Running centroid variation for primates (Parth)
- Everyone do documentation (centroid, limb ratio, dino, downstream tasks, labeling strategy…)

VIP

PURDUE UNIVERSITY

DINOv2 + Stable Diffusion Features

Motivation: DINOv2(sparse meaningful points) and SD(general features) features seem to complement each other(https://arxiv.org/abs/2305.15347) and so should give more descriptive features when combined.

Procedure:

- Cropped input images at their bounding box then scaled(preserved aspect ratio and added padding) to 840x840(size used in the linked paper)
- Expanded stable diffusion features from 2([batch, feature_dim]) -> 3([batch, dino_num_patches, sd_feature_dim]), where replicated the original feature_dim sized features over each patch
- Concatenated stable diffusion and dinov2 features on the feature dimension ([batch, dino_num_patches, sd_feature_dim + dinov2_feature_dim])
- Calculated features and evaluated species similarity based on only side-view images of each species, and all images from AP-10k of those species, and generated lists of similar species based the features from each of those datasets.
- Trained RTMPose models on the top 10 similar species from each dataset and based on the cosine similarity metric and the KNN metric

Possible Improvements:

- Test other manners of fusing the SD and DINOv2 features
- Better utilization of PCA(reference how linked paper uses PCA)
- Other Metrics(human visual similarity for proof of theory)?

VIP

PURDUE UNIVERSITY

Similar Species - DINOv2 + Stable Diffusion Features

Before Cosine
Side-view

```
antelope: 1.0000
deer: 0.9165
fox: 0.8886
dog: 0.8702
weasel: 0.8641
sheep: 0.8583
mouse: 0.8569
raccoon: 0.8528
spider monkey: 0.8495
cheetah: 0.8455
squirrel: 0.8422
argali sheep: 0.8421
rabbit: 0.8417
hamster: 0.8400
cow: 0.8390
moose: 0.8386
rat: 0.8359
bobcat: 0.8325
giraffe: 0.8318
wolf: 0.8316
brown bear: 0.8286
polar bear: 0.8235
bison: 0.8234
otter: 0.8225
monkey: 0.8186
snow leopard: 0.8153
hippo: 0.8132
pig: 0.8048
cat: 0.8037
beaver: 0.8036
horse: 0.7979
panda: 0.7909
skunk: 0.7833
rhino: 0.7786
buffalo: 0.7689
tiger: 0.7663
leopard: 0.7577
noisy night monkey: 0.7563
jaguar: 0.7536
elephant: 0.7532
lion: 0.7530
chimpanzee: 0.7381
marmot: 0.7168
alouatta: 0.6880
zebra: 0.6616
panther: 0.6576
black bear: 0.6222
king cheetah: 0.6012
gorilla: 0.4952
```

After Cosine
Side-view

```
antelope: 1.0000
deer: 0.9020
giraffe: 0.8307
cheetah: 0.8216
argali sheep: 0.8103
moose: 0.8085
fox: 0.8077
buffalo: 0.8063
sheep: 0.8037
zebra: 0.8018
rabbit: 0.7990
leopard: 0.7968
bobcat: 0.7955
cow: 0.7949
wolf: 0.7895
bison: 0.7894
spider monkey: 0.7894
monkey: 0.7888
weasel: 0.7875
elephant: 0.7798
lion: 0.7752
tiger: 0.7737
squirrel: 0.7701
dog: 0.7674
rhino: 0.7647
hippo: 0.7630
pig: 0.7556
jaguar: 0.7556
brown bear: 0.7549
raccoon: 0.7487
mouse: 0.7487
polar bear: 0.7468
horse: 0.7461
snow leopard: 0.7381
otter: 0.7373
beaver: 0.7351
king cheetah: 0.7341
panda: 0.7317
rat: 0.7257
skunk: 0.7244
marmot: 0.7227
noisy night monkey: 0.7222
cat: 0.7159
chimpanzee: 0.7121
alouatta: 0.7060
panther: 0.6988
hamster: 0.6717
black bear: 0.6236
gorilla: 0.6136
```

Before KNN
Side-view

```
weasel: 15
deer: 14
fox: 12
moose: 10
rabbit: 10
mouse: 7
sheep: 5
raccoon: 5
cheetah: 4
giraffe: 4
squirrel: 3
dog: 3
rat: 2
bobcat: 2
polar bear: 2
otter: 2
hippo: 1
horse: 1
cow: 1
cat: 1
argali sheep: 1
brown bear: 1
skunk: 1
spider monkey: 1
panda: 1
beaver: 1
```

After Cosine
Side-view

```
deer: 35
giraffe: 27
moose: 17
bison: 8
rabbit: 7
cheetah: 4
argali sheep: 4
leopard: 2
sheep: 2
zebra: 2
bobcat: 1
fox: 1
```

Cosine
Full AP-10k

```
antelope: 1.0000
deer: 0.9020
giraffe: 0.8307
cheetah: 0.8216
argali sheep: 0.8103
moose: 0.8085
fox: 0.8077
buffalo: 0.8063
sheep: 0.8037
zebra: 0.8018
rabbit: 0.7990
leopard: 0.7968
bobcat: 0.7955
cow: 0.7949
wolf: 0.7895
bison: 0.7894
spider monkey: 0.7894
monkey: 0.7888
weasel: 0.7875
elephant: 0.7798
lion: 0.7752
tiger: 0.7737
squirrel: 0.7701
dog: 0.7674
rhino: 0.7647
hippo: 0.7630
pig: 0.7556
jaguar: 0.7556
brown bear: 0.7549
raccoon: 0.7487
mouse: 0.7487
polar bear: 0.7468
horse: 0.7461
snow leopard: 0.7381
otter: 0.7373
beaver: 0.7351
king cheetah: 0.7341
panda: 0.7317
rat: 0.7257
skunk: 0.7244
marmot: 0.7227
noisy night monkey: 0.7222
cat: 0.7159
chimpanzee: 0.7121
alouatta: 0.7060
panther: 0.6988
hamster: 0.6717
black bear: 0.6236
gorilla: 0.6136
```

KNN Full AP-
10k

```
deer: 35
giraffe: 27
moose: 17
bison: 8
rabbit: 7
cheetah: 4
argali sheep: 4
sheep: 2
zebra: 2
leopard: 2
fox: 1
bobcat: 1
```

VIP

PURDUE UNIVERSITY

Average Precision - DINOv2 + Stable Diffusion Features

	Full AP-10k redistributed	Top 10 Dino Full AP-10k Cosine Similarity	Top 10 Dino Full AP-10k KNN	Top 10 Dino Side View Cosine Similarity	Top 10 Dino Side View KNN
coco/AP:	0.818	0.778	0.756	0.755	0.729
coco/AP .5:	0.969	0.964	0.962	0.935	0.931
coco/AP .75	0.884	0.842	0.811	0.802	0.788
coco/AP (M):	0.799	0.605	0.736	0.684	0.607
coco/AP (L):	0.817	0.788	0.761	0.761	0.737

VIP

PURDUE UNIVERSITY

Average Recall - DINOv2 + Stable Diffusion Features

	Full AP-10k redistributed	Top 10 Dino Full AP-10k Cosine Similarity	Top 10 Dino Full AP-10k KNN	Top 10 Dino Side View Cosine Similarity	Top 10 Dino Side View KNN
coco/AR:	0.834	0.801	0.785	0.780	0.762
coco/AR .5:	0.975	0.971	0.968	0.943	0.943
coco/AR .75	0.892	0.856	0.836	0.823	0.815
coco/AR (M):	0.817	0.609	0.758	0.687	0.610
coco/AR (L):	0.834	0.811	0.789	0.786	0.770

Primate Experiments with Centroid Metric

- To see if the centroid metric can be extended to other classes, it is worth comparing the centroid metric to ap-10k for other species such as some within the class of primates
- Picked chimpanzee as the species to run centroid metric.
- Motivation behind this choice is that, chimpanzees have a lot of shared human traits, so the movements/body structure will be significantly different from that of an antelope.

Primate Experiments Procedure

- Separated chimpanzees from AP-10k into a testing set
  - Chimpanzee-only test set: 200 images
- Redistributed AP-10K without chimpanzee with train/val/test -> train/val
  - AP10K (no chimpanzee) train set: 7852 images
  - AP10K (no chimpanzee) val set: 1963 images
- Found top 10 species with centroid metric for chimpanzees and created dataset
  - Top 10 total images: 1674, Top 10 train set size: 1340, Top 10 val set size: 334
  - Top 10 list on next slide
- Downsized AP-10k that was redistributed to:
  - Original train images: 7852, Downsized train images: 1340
  - Original val images: 1963, Downsized val images: 334
  - Total downsized images: 1674

Top & Bottom 10 closest species to chimpanzee's

Top 10 Species Most Similar to chimpanzee:	
uakari	0.997240
gorilla	0.996809
alouatta	0.994453
spider monkey	0.992563
noisy night monkey	0.991660
monkey	0.991623
panda	0.990544
hamster	0.982954
polar bear	0.982631
brown bear	0.981697

Bottom 10 Species Most Similar to chimpanzee:	
rhino	0.959396
squirrel	0.957245
zebra	0.956389
hippo	0.951456
horse	0.950019
argali sheep	0.949485
bison	0.944096
antelope	0.940325
rabbit	0.939396
giraffe	0.932889

Interesting note: All 6 primates within AP-10k are within the top 6 of the list

VIP

PURDUE UNIVERSITY

Model Testing Results

AP - Average Precision, AR - Average Recall on 200 chimpanzee images

	Full AP-10k	Downsized AP-10k	Top 10 Centroid
coco/AP:	0.641	0.546	0.693
coco/AP .5:	0.903	0.867	0.949
coco/AP .75	0.701	0.596	0.766
coco/AP (M):	0.242	0.143	0.542
coco/AP (L):	0.670	0.573	0.705

	Full AP-10k	Downsized AP-10k	Top 10 Centroid
coco/AR:	0.682	0.587	0.729
coco/AR .5:	0.912	0.882	0.958
coco/AR .75	0.737	0.634	0.798
coco/AR (M):	0.332	0.263	0.595
coco/AR (L):	0.709	0.613	0.740

VIP

PURDUE UNIVERSITY

Centroid Variation (Side-view vs Full AP10k)

- To create the top ten most similar species to antelopes using the centroid variation method, we try two different datasets for comparisons
- 1. Full AP10k images to compute centroid variation per species
- 2. Side-view images (to see if keeping the centroid in a relatively static position on the antelope can increase information gained about centroid variation

Top 10 antelopes using side view

horse	0.996515
moose	0.992836
sheep	0.992651
deer	0.991479
fox	0.988967
cow	0.988172
skunk	0.987897
brown bear	0.987722
cheetah	0.986280
argali sheep	0.985105

Top 10 antelopes using AP10k

argali sheep	0.998723
horse	0.997937
deer	0.994156
zebra	0.992842
moose	0.992537
giraffe	0.991767
king cheetah	0.989388
sheep	0.987457
bison	0.987215
fox	0.986510

Key differences: ranking of argali sheep, addition of skunk and brown bear in side-view top 10

VIP

PURDUE UNIVERSITY

Results of training with top 10 species (centroid variation)

Both datasets were created to be the same size (1638 images with an 80/20 test/val split)

	Top 10 Centroid (full AP10k)	Top 10 Centroid (side-view images)
coco/AP:	0.821	0.817
coco/AP .5:	0.978	0.967
coco/AP .75	0.904	0.894
coco/AP (M):	0.721	0.772
coco/AP (L):	0.822	0.817

	Top 10 Centroid (full AP10k)	Top 10 Centroid (side-view images)
coco/AR:	0.843	0.836
coco/AR .5:	0.980	0.975
coco/AR .75	0.917	0.907
coco/AR (M):	0.767	0.783
coco/AR (L):	0.845	0.837

VIP

PURDUE UNIVERSITY

Limb Ratio (Side View vs. Full AP10k)

- Previously we only used the side view images to calculate the limb ratio and get the top 10 most similar species to the antelope
- Now we use the full AP10k images to calculate the top 10 most similar species.

Top 10 by using Side View

```
Top Ten Species Most Similiar to Antelope
argali sheep: 0.9954
horse: 0.9886
moose: 0.9812
dog: 0.9799
zebra: 0.9795
deer: 0.9777
sheep: 0.9776
cow: 0.9726
fox: 0.9694
buffalo: 0.9692
```

Top 10 by using full Ap-10k

```
Top Ten Species Most Similiar to Antelope
deer: 0.9938
argali sheep: 0.9917
horse: 0.991
raccoon: 0.9896
brown bear: 0.9896
zebra: 0.989
moose: 0.9881
panda: 0.9851
polar bear: 0.9842
skunk: 0.9841
```

VIP

PURDUE UNIVERSITY

Testing Result (limb ratio)

Both dataset have 1811 images and 80:20 training (1449 train 362 val)

	Top 10 Limb ratio (full AP10k)	Top 10 Limb ratio (side-view images)
coco/AP:	0.789	0.794
coco/AP .5:	0.968	0.957
coco/AP .75	0.866	0.873
coco/AP (M):	0.740	0.722
coco/AP (L):	0.788	0.794

	Top 10 Limb ratio (full AP10k)	Top 10 Limb ratio (side-view images)
coco/AR:	0.810	0.815
coco/AR .5:	0.970	0.960
coco/AR .75	0.877	0.882
coco/AR (M):	0.767	0.733
coco/AR (L):	0.812	0.818

VIP

PURDUE UNIVERSITY

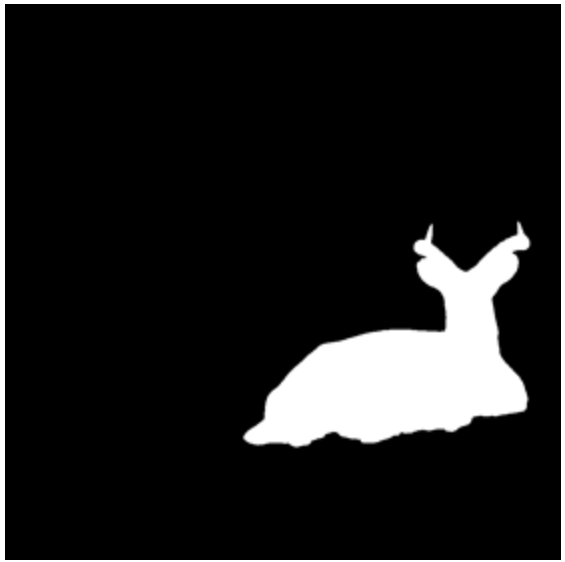
Next Steps

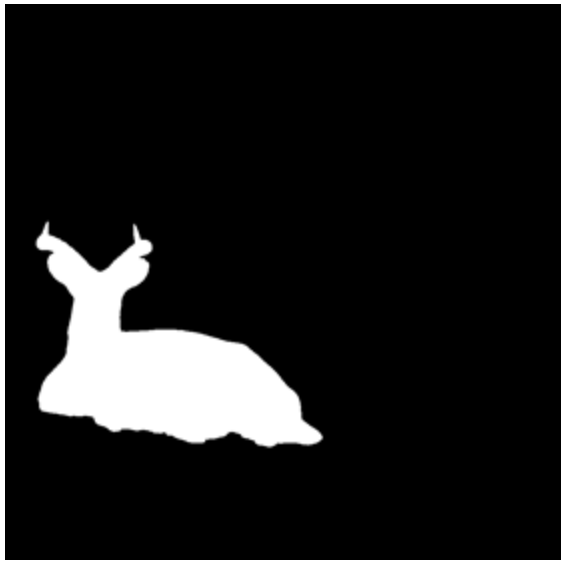
- Continue working on AP-10K labeling effort
- Continue species similarity experiments

Label Studio

Semantic Correspondence Matching

- Stable Diffusion and DINOv2 features extracted for all animal images





Personal Progress

Medha

- Created top 10 list for centroid variation using side-view images
- Trained and tested RTMPose for both top-10 side-view based dataset and a reduced top-10 AP10k based dataset for comparison
- Labeled 40 images with visible keypoint definition

Shaan

- Set up Label Studio with antelope images and AP10k labels pre-loaded
- Ran test-time pose alignment on Antelope images

Parth

- Created top 10 centroid list for chimpanzees for primate experiments
- Trained & Tested this top 10 centroid list for chimpanzees, AP-10k without chimpanzees and downsized ap-10k without chimpanzees (to match the top 10 list length)
- Created all datasets and scripts required for training & testing
- Labeled all AP-10k Antelopes with required keypoints for visible definition